Dimensionality Reduction Techniques Applied to the Design of Hypersonic Aerial Systems

Kenneth H. Decker*, Henry D. Schwartz†, and Dimitri N. Mavris‡
Aerospace Systems Design Laboratory, Georgia Institute of Technology, Atlanta, GA, 30332 USA

This study presents a methodology for the use of parametric Reduced Order Modeling (ROM) techniques to generate predictive models of hypersonic aerodynamic flow fields. The goal of this study is to synthesize a methodology for the development of these field surrogate models using techniques and procedures from the literature. This methodology is applied to two analytical test problems and one CFD application to demonstrate the functionality of the methodology and quantify the performance of models generated using various ROM techniques. The models compared in this study were generated using Proper Orthogonal Decomposition (POD) as a representative linear dimensionality reduction method, along with ISOMAP and Locally Linear Embedding (LLE) as representative Nonlinear Dimensionality Reduction (NLDR) methods. Based on the results of study, it is observed that the NLDR-based ROMs provide better predictions in the regions of fields near shocks, while linear methods are found to outperform non-linear methods when predicting steady-state behaviors far from shocks. Furthermore, nonlinear ROMs generated models of lower dimension than their linear counterparts, which resulted in significantly lower evaluation cost and could have significant ramifications if these predictive models are applied to coupled systems analyses.

NOMENCLATURE

a	=	Step size
c	=	Cost term
d	=	ROM dimension
d_{ij}	=	Pairwise distance operator
f, g	=	Generic functions
i, j, k, l	=	Generic indices
k	=	Number of nearest neighbors
M	=	Mach number
m	=	Number of training snapshots
n	=	Dimension of the field
$\mathbf{p}$	=	Vector of design parameters
p	=	Pressure
v	=	Number of validation snapshots
w	=	Weight coefficients
$\mathbf{W}$	=	Snapshot of field data
$\mathbf{y}$	=	Low-dimensional representation of field
α	=	Angle of attack
β	=	Shock angle
γ	=	Model tuning parameter, ratio of specific heats
δ	=	Turning angle
ϵ	=	Cost function weight
θ	=	Generic angle
σ	=	Singular value
τ	=	Shock vicinity threshold
ϕ	=	POD mode

*Ph.D. Candidate, School of Aerospace Engineering (ASDL), Georgia Tech, Student AIAA Member
†Research Engineer II, School of Aerospace Engineering (ASDL), Georgia Tech, AIAA Member
‡S.P. Langley Distinguished Regents Professor and Director of ASDL, Georgia Tech, AIAA Fellow

CFD	=	Computational Fluid Dynamics
DoE	=	Design of Experiments
DoF	=	Degree of Freedom
DR	=	Dimensionality Reduction
FOM	=	Full Order Model
FUN3D	=	Fully Unstructured Navier-Stokes 3-Dimensional Solver
LLE	=	Locally Linear Embedding
MDAO	=	Multidisciplinary Design Analysis and Optimization
NLDR	=	Non-Linear Dimensionality Reduction
PCA	=	Principal Component Analysis
POD+I	=	Proper Orthogonal Decomposition plus Interpolation
RBF	=	Radial Basis Function
RIC	=	Relative Information Content
ROM	=	Reduced Order Model
SVD	=	Singular Value Decomposition
TPS	=	Thin-Plate Spline

I. Introduction

The demands for the next generation of hypersonic vehicles are growing increasingly complex and diverse. Entities from the defense, commercial, and civil domains have expressed interest in the development of novel hypersonic aerial systems to be designed for a variety of tasks. Traditional conceptual design methods are dependent upon the use of historical data to perform initial conceptual sizing and synthesis [1]. However, when designing novel concepts, it can be difficult, or even impossible, to evaluate design performance based on historical data because the necessary data likely do not exist [2]. Thus, the design of novel concepts for which historical data are unavailable requires novel processes to generate the information needed to carry out systems engineering tasks, particularly during the early design phases [2]. Today, designers have access to well-developed physics-based simulation tools that enable realistic estimations of system behavior and performance that can be made at a fraction of the time and cost of physical experiments [3].

Engineers must be able to evaluate system-level performance to determine if multidisciplinary designs are feasible, if disciplines can be integrated, and if the integrated system is capable of meeting the requirements posed by stakeholders. Hypersonic systems are unique because they contain unique disciplines that interact in ways that are not seen in other classes of aircraft. Hypersonic aerial systems are, by necessity, highly-integrated systems that are typically defined by complex interactions that arise from the tight coupling between disciplines [4]. Typically, when analyzing complex systems, the fundamental relationships between inputs and outputs are unknown at the system-level and cannot be inferred from analyzing individual disciplines or derived from first principles [5, 6]. The design of systems of this nature requires large amounts of information to identify a system design that is both feasible and optimal. Consequently, when designing novel aerial systems, designers may need to evaluate thousands or even millions of performance simulations [7].

A central theme in physics-based modeling is the trade-off between model fidelity and computational cost. High-fidelity tools are desirable because they generally capture a large amount of the underlying physics of the problem and produce reasonably accurate simulations, but the labor and computational cost that is necessary to implement high-fidelity tools impose practical constraints on their use. Meanwhile, low-fidelity tools operate at a much lower computational cost, but are limited in their accuracy and are often only able to capture general trends and sensitivities [8]. For hypersonic aerial systems, quantities such as pressure distributions, forces, and moments can be reasonably predicted using low-fidelity methods; however, metrics driven by phenomena such as off-body flow features, shock-viscous interactions, structural dynamics, and non-ideal thermochemistry require high-fidelity analyses to capture at a reasonable level of uncertainty [9–13]. To be successful, system models must identify all potential relevant physical and functional impacts [14]. For this reason, a current goal of systems engineering is to increase the fidelity of the analyses that designers can implement in the early design processes to enable more effective decision support [8].

A common approach to improving the computational efficiency of coupled system analyses is the implementation of disciplinary surrogate models. To obtain the necessary information, designers typically perform a Design of Experiments (DoE), in which design cases are selectively evaluated with the goal of maximizing information obtained about a design

space while minimizing the necessary input cost [15, 16]. Surrogate models are then trained using statistical methods to approximate a complex analysis model in a way that preserves the accuracy of the original analysis over the sampled range of parameters in a form that is relatively inexpensive to evaluate. The goal is to evaluate a manageable number of designs and use these data to train and validate a surrogate model prior to performing a coupled analysis (referred to as the "offline" phase). Then, the validated surrogate models replace the high-fidelity analysis during system optimization (referred to as the "online" phase) when the surrogate model is evaluated many times over the course of an optimization. Since the computational cost of evaluating the surrogate model is typically small with respect to the original high-fidelity analysis, computational savings is realized over the complete process [17]

Surrogate models that map scalar inputs to scalar outputs have been commonly applied in previous studies and methods for their implementation have been well developed. However, scalar data are inherently limited because they cannot describe quantities that vary over the body in detail. This is especially relevant for hypersonic systems, whose flow fields are characterized by strong shocks and expansions. These strong, discontinuous variations in properties over a vehicle surface can have significant impacts on thermal protection system design, structural sizing constraints, aeroelastic performance, scramjet inlet design, stability and control performance, and many other critical design aspects [18–20]. Therefore, it is critical that designers have access to reliable field data early in the design process to inform decision-making and produce effective, optimal systems.

Field surrogate modeling, or Reduced Order Modeling (ROM), has been identified as a key enabler for Multi-disciplinary Design, Analysis, and Optimization (MDAO) of integrated hypersonic aerial systems [17]. A challenge associated with generating field surrogate models is that the underlying form of the model that best represents the observed behavior of the field is unknown a priori. With some scalar-to-scalar surrogate models, analysts can assume an underlying form of the model (e. g. response surface equations often assume the training data can be approximated as a second order polynomial) [21]. However, this is not practical for field data due to their size and complexity. Instead, Dimensionality Reduction (DR) techniques can be applied to identify a low-dimensional representation of the field data that captures the relevant observed behavior that is then used to develop predictive ROMs [22]. DR techniques can be broadly classified as linear or nonlinear, both of which are discussed in detail in a following section of this paper. Linear DR techniques, such as Proper Orthogonal Decomposition (POD), have been applied in aerodynamics for decades [23–25]. However, Nonlinear Dimensionality Reduction (NLDR) techniques have gotten more recent attention due to their ability to represent nonlinear field features, such as shocks [26, 27]. In either case, the predictive ROMs generated using these low-dimensional bases enable high-fidelity field data to be approximated in a compact, computationally efficient manner.

The objective of this study is to investigate the use of DR techniques to generate surrogate models of hypersonic flow fields. Linear and nonlinear DR techniques will be investigated to determine which methods produce more accurate ROMs of fields that contain discontinuous features, such as shocks. This study will examine the sensitivity of surrogate model accuracy to factors such as the number of training cases, field discretization topology, and method-specific tuning parameters in an attempt to establish a repeatable and robust process for model generation. The mathematical formulations of the proposed ROM techniques are discussed in the following section. Next, the details of the computational methods used to implement these field surrogate models are discussed. Then, test problems are derived and analyzed to assess the ability of the proposed methodology to create accurate ROMs. Finally, a the results of these test problems are discussed and fundamental conclusions are formulated.

II. Reduced Order Modeling

Two broad classifications of ROMs have been developed in the literature: intrusive ROMs and non-intrusive ROMs. Intrusive ROMs generally use sample data along with knowledge of the underlying governing equations of an analysis model to obtain a reduced order projection [28]. While rooting ROMs in the underlying governing equations can be attractive and improve model accuracy, these methods risk introducing instabilities into models as the underlying governing equations grow in dimension and complexity [29]. Furthermore, these methods often require modifications to be made to underlying source code and are thus not conducive to use with "black box" models [30]. Non-intrusive solve these issues by only requiring input and output data to fit models. Consequently, non-intrusive models are more widely applicable and can handle "black box" analysis codes, which make them ideal candidates for highly multidisciplinary problems [30, 31]. For these reasons, non-intrusive models are chosen for this study.

Non-intrusive ROM techniques can broadly be grouped into two categories: linear models and nonlinear models. Linear and nonlinear ROMs vary significantly in their properties and underlying assumptions, which can have drive the ability of these models to accurately capture the underlying behavior of a given data set and make predictions. The

details of these properties, and their relevance to hypersonic field surrogate modeling, are discussed in the following sections.

A. Linear Dimensionality Reduction Methods

Perhaps the most commonly used ROM technique is Proper Orthogonal Decomposition (POD), which is also referred to in the literature as Principal Component Analysis (PCA) or Karhunen-Loève Decomposition [32]. POD seeks to project discretized, high-dimensional data into a low-dimensional linear subspace based on sample data, or "snapshots", obtained at various points in the design space. In POD, the basis functions, or "modes", are identified using Singular Value Decomposition (SVD). Consider a matrix W that is composed of m snapshots of a field of dimension n obtained from numerical simulation such that $W = \left[\mathbf{W}^1, ..., \mathbf{W}^m\right] \in \mathbb{R}^{nxm}$. SVD can be performed on W such that:

$$W = U\Sigma V^T \tag{1}$$

Where $U \in \mathbb{R}^{nxn}$, $V \in \mathbb{R}^{mxm}$, and $\Sigma = diag(\sigma_1, ..., \sigma_r,) \in \mathbb{R}^{nxm}$ such that $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r \geq 0$ and $r = min(n, m)$; the squared values of $\sigma_1^2, ..., \sigma_p^2$ correspond to the non-zero eigenvalues of W^TW and WW^T whose corresponding eigenvectors form the columns of V and U respectively. If the matrix W has rank r where $r \leq m \leq n$, then the first r columns of the matrix U exhibit the property that:

$$span\{\mathbf{U}^1, ..., \mathbf{U}^r\} = span\{\mathbf{W}^1, ..., \mathbf{W}^m\} \tag{2}$$

Where the set $\{\mathbf{W}^1, ..., \mathbf{W}^m\}$ denotes the columns of W and the set $\{\mathbf{U}^1, ..., \mathbf{U}^r\}$ denotes the first r columns of U. A lower rank approximation of the original data can be obtained by computing a matrix $X \in \mathbb{R}^{nxm}$ with a rank of d such that $d < r$ [32]. To do this, Σ (and the corresponding columns of U and V) is arranged such that its entries are in descending order. The reduced rank matrix, X is then obtained by preserving the first d columns of U, which correspond to the d largest singular values in Σ. According to the Eckhart-Young Theorem, if $X = U\Sigma_k V^T = \sum_{i=1}^{d} \sigma_i \mathbf{U}^i \left(\mathbf{V}^i\right)^T$ where Σ_d is a modified version of the original Σ matrix where $\sigma_{d+1} = \sigma_{d+2} = ... = \sigma_r = 0$, then:

$$\min_{rank(X) \leq d} ||W - X||_2^2 = \sum_{i=d+1}^{r} \sigma_i^2 \tag{3}$$

$$\min_{rank(X) \leq d} ||W - X||_F^2 = \sigma_{d+1} \tag{4}$$

Where $|| \cdot ||_F$ denotes the Frobenius norm. A consequence of Equations 3 and 4 is that the basis $\{\mathbf{U}^1, ...\mathbf{U}^d\}$ leads to the best rank-d approximation of W according to either the 2-norm or Frobenius norm. A detailed derivation of Equations 3 and 4 as well as the proof of this consequence is presented in Reference [32]. The modes obtained from POD are uncorrelated (orthogonal) and capture as much variation of the original snapshots as possible, making the reduced basis the optimal linear subspace to represent the original data [33]. Since the rank-d approximation error is $\sum_{i=d+1}^{r} \sigma_i^2$, it is possible to define the Relative Information Content (RIC) of including d basis vectors in an approximation as follows [32]:

$$RIC(d) = \frac{\sum_{i=1}^{d} \sigma_i^2}{\sum_{i=1}^{r} \sigma_i^2} \tag{5}$$

Typically, the number of basis vectors d that are preserved in a POD approximation is not known a priori. Instead, the analyst selects a desired information content $\delta \in [0, 1]$ to preserve, then selects d basis vectors such that $RIC(d) \geq \delta$, with values of $\delta \geq .99$ seen commonly in application [27, 34].

Each snapshot $\mathbf{W}^i$ is generated based on some known combination of parameters $\mathbf{p}^i$. A property of SVD is that the resulting set of basis functions are mutually orthogonal. Thus, for any of the known snapshots, it is possible to determine the coefficient values on each of the basis functions using inner products [32]. If the coefficient values are smooth functions of $\mathbf{p}$, then interpolation can be used to determine the POD coefficients for interior values of $\mathbf{p}$, allowing the flow field predictions at unsampled values of $\mathbf{p}$ [23]. The application of interpolation to obtain the POD mode coefficients is referred to as "POD with Interpolation", or "POD+I". The mathematical formulation of the resulting predictive field surrogate model for parametric problems discretized in space is given by the following equation:

$$\mathbf{W}(\mathbf{x}; \mathbf{p}) \approx \mu(\mathbf{x}) + \sum_{i=1}^{d} a_i(\mathbf{p})\phi_i(\mathbf{x}) \tag{6}$$

Where $\mathbf{W}(\mathbf{x};\mathbf{p})$ is the predicted field, $\mathbf{p}$ is a set of parameters of interest other than space and time. The field is approximated by its mean value in space, $\mu(\mathbf{x})$, and a sum of d basis functions in space, $\phi_i(\mathbf{x})$, with parameterized weight coefficients, $a_i(\mathbf{p})$ [23, 33].

The linear nature of methods like POD makes them intuitive and simple to implement. POD has been applied to aerodynamic flow fields across the full spectrum of flow regimes [7, 27, 32, 35–39]. While POD has been successfully applied to numerous applications, it is inherently limited by the assumption that the underlying behavior can be projected into a linear subspace spanned by the original data [7]. Features that are exhibited in hypersonic flows, such as shocks, can result in highly nonlinear behavior and thus violate the underlying assumptions of POD, potentially limiting its capabilities.

B. Nonlinear Dimensionality Reduction Methods

When the underlying behavior of a problem is sufficiently nonlinear, NLDR algorithms can potentially outperform the previously discussed linear methods when identifying critical underlying behaviors in a given set of data [40]. Manifold Learning algorithms are a family of NLDR methods that are commonly used to identify a low-dimensional nonlinear manifold that accounts for the observed variability in the high-dimensional data [41]. The fundamental goal of manifold learning algorithms is to solve the embedding problem, which is stated as follows by Bernstein and Kuleshov:

> For a given data set $W = \{\mathbf{W}^1,...,\mathbf{W}^m\} \subset \mathcal{W} \subset \mathbb{R}^n$ sampled from an unknown submanifold $\mathcal{W}$ in Euclidean space $\mathbb{R}^n$ with intrinsic dimensionality $dim\mathcal{W} = d < n$ (usually $d << n$ is to be expected), find an embedding mapping
>
> $$\mathbf{h} : \mathcal{W} \mapsto \mathbb{R}^d, \quad \mathbf{h}(\mathbf{W}^i) = \mathbf{y}^i \in \mathbb{R}^d \tag{7}$$
>
> which preserves the geometry of the data set W as much as possible, so that the low dimensional data set $Y = \{\mathbf{y}^1,...,\mathbf{y}^m\}$ is a good representation of the high-dimensional data set W [42].

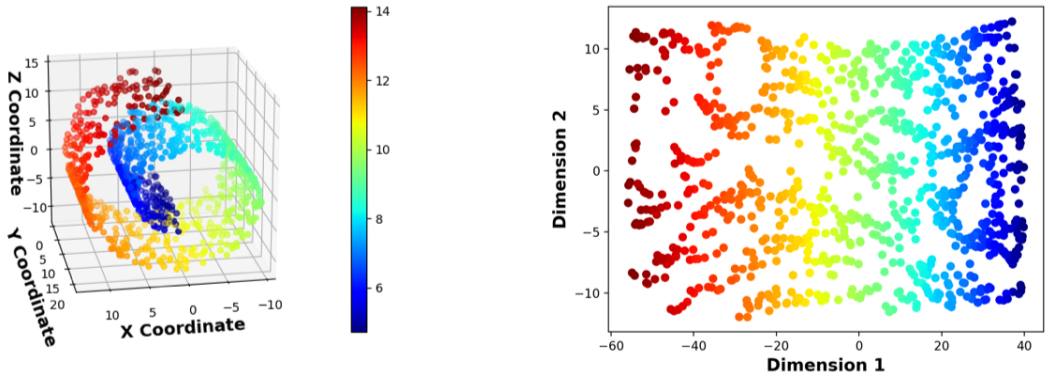
Thus, the objective of manifold learning methods is to identify the low-dimensional structures in high-dimensional data such that the behavior of the high-dimensional data can be efficiently analyzed [27]. There are many different types of manifold algorithms that have been developed in the literature. Each of these methods seeks to solve the same fundamental problem, but primarily differ in the metrics used to obtain the "best" low-dimensional embedding through optimization [39]. Given a set of snapshots for a numerical simulation, all manifold learning algorithms broadly follow the following steps, according to a survey by Izenman [22]:

1) Construct a graph with the data points as vertices using neighborhood information around each point
2) Transform graph into a form that is suitable for the next step using a procedure unique to the algorithm
3) Perform a spectral embedding by solving an eigenproblem

Each algorithm yields an embedding $Y = \{\mathbf{y}^1,...,\mathbf{y}^m\}$ where $\mathbf{y}^i \in \mathbb{R}^d$ and is the low-dimensional representation of the full order snapshot $\mathbf{W}^i \in \mathbb{R}^n$ [27]. A simple illustration of this is shown in Figure 1, which is obtained by applying manifold learning (ISOMAP in this example) to a "Swiss roll" test problem. This example shows data that appear 3-D, as shown in Figure 1a. However, the data exhibit intrinsically 2-D behavior, which is successfully recovered by the manifold learning algorithm in Figure 1b.

In general, manifold learning problems are not well-posed because the intrinsic dimensionality, d, of the problem is not necessarily known a priori. However, when applied to a data set generated by varying a set of parameters, $\mathbf{p}$, whose dimension is known, it is appropriate to assume that the intrinsic dimensionality of the problem is equal to the number of parameters in $\mathbf{p}$ [27]. Manifold algorithms can either be local or global in nature, where local methods seek to preserve relationships of points within local regions of the data space while global methods seek to preserve relationships across all points [22]. Local methods tend to be more computationally efficient and better-conditioned while global methods can be more accurate when suitable [22]. For this study, ISOMAP is identified as a representative global method while LLE is identified as representative local method. Many other methods have been developed, but these two are used as representative methods in this study due to their maturity. Detailed derivations for these methods are presented in References [40] and [43] respectively.

Similar to POD+I, interpolation can then be applied to the low-dimensional representation to approximate the coordinates of unsampled points in this low-dimensional space [26]. If the mapping $\mathbf{p} \mapsto \mathbf{W}(\mathbf{p})$ is smooth and the embedding is known, it can be assumed that the mapping $\mathbf{p} \mapsto \mathbf{y}^i$ is also smooth [26]. Thus, an interpolate model can be fit to estimate $\mathbf{y}(\mathbf{p})$ for a set of parameters $\mathbf{p}$ within the sampled space of the original snapshots. Then, the high-dimensional reconstruction of the field can be obtained based on the estimation of low-dimensional parameters according to various back-mapping techniques [26].



(a) Original Data **(b) Low-Dimensional Representation**

Fig. 1 Application of Manifold Learning to Swiss Roll Test Problem

A back-mapping has been applied in previous studies based on the assumption that the data can be treated as locally linear [26, 44]. Consider the prediction of a field at a set of parameters $\mathbf{p}^*$. If the values of $\mathbf{p}$ and their corresponding low-dimensional manifold coordinates, $\mathbf{y}$, are known at the set of snapshots, then an interpolate method can be used to map the prediction point $\mathbf{p}^*$ to a prediction point on the manifold, $\mathbf{y}^*$. A k-nearest-neighbor method can be used to identify the k snapshots whose low-dimensional projections, $\mathbf{y}_j$, are closest to $\mathbf{y}^*$. By assuming that the geometry of the recovered manifold is locally linear, it is possible to compute a set of weight coefficients, w_j that approximate $\mathbf{y}^*$ as a linear combination of its k nearest neighbors along the manifold, such that $\mathbf{y}^* \approx \sum_{j=1}^{k} w_j \mathbf{y}_j$ [26, 43]. The values of the weight coefficients, w_j, can be computed by solving a least squares optimization subject to the constraint that $\sum_{j=1}^{k} w_j = 1$, with the constraint ensuring that the approximation is invariant under linear translation [26]. However, if the number of nearest neighbors is greater than the dimension of the manifold, then the constrained least squares problem does not have a unique solution. To address this issue, a method has been presented by Franz based on work by Saul and Roweis which adds the following terms to the optimization problem [27, 43]:

$$\|w\|_c^2 = \sum_{j=1}^{k} c_j w_j^2 \tag{8}$$

$$c_j = \epsilon \left(\frac{\|\mathbf{y}^* - \mathbf{y}^j\|_2}{max_i \|\mathbf{y}^* - \mathbf{y}^i\|_2} \right)^{\gamma} \tag{9}$$

Where ϵ and γ are user defined tuning parameters chosen such that $0 < \epsilon \ll 1$ and $1 < \gamma \in \mathbb{N}$. The quantity c_j is effectively a penalty term applied to the j^{th} snapshot that works to both regularize the optimization and reduce the contribution that snapshots far from $\mathbf{y}^*$ make to its reconstruction. The result is the following optimization problem, which has a unique solution and results in points that are closer to $\mathbf{y}^*$ having a larger contribution to its reconstruction:

$$\min_{w} \|\mathbf{y}^* - \sum_{j=1}^{k} w_j \mathbf{y}^j\|_2^2 + \|w\|_c^2 \tag{10}$$

$$s.t. \sum_{j=1}^{k} w_j = 1 \tag{11}$$

The weight coefficients obtained by solving Equations 10-11 can be used to reconstruct the high dimensional data as a linear combination of the corresponding high-dimensional snapshots, according to the equation:

$$\mathbf{W}^* = \sum_{j=1}^{k} w_j^* \mathbf{W}^j \tag{12}$$

Where $\mathbf{W}^*$ is the reconstructed high-dimensional data at $\mathbf{p}^*$, w_j^* are the weight coefficients that solve Equations 10-11, and $\mathbf{W}_j$ are the high-dimensional snapshots that correspond to the k nearest neighbors of $\mathbf{y}^*$in the recovered manifold. The result of this process is a parametric, interpolation-based ROM that allows the user to predict field data at an arbitrary point $\mathbf{p}^*$ that lies within the domain sampled by the training data.

NLDR candidates do not presume any particular form for underlying behavior of the data, and the assumption of local linearity made by the back-mapping procedure used in this study is less constraining than the assumption of global linearity made by methods like POD. Consequently, NLDR algorithms have the potential to be applied to a wider range of problems and could be promising candidates for bringing high-fidelity information into MDAO applications at a feasible computational cost, even when the data being modeled is highly nonlinear.

III. Methodology

The previous section presents mathematical formulations for the DR and back-mapping techniques that are the basis for developing predictive field surrogate models. A primary goal of this study is to implement these concepts into a repeatable methodology that can be used to generate predictive field surrogate models for generic aerospace problems. Once implemented, this methodology can then be applied to various applications and test problems to characterize its performance. The proposed methodology consists of four fundamental procedures:

1) Data generation through evaluation of the Full Order Model (FOM)
2) Dimensionality reduction and projection into a reduced order representation
3) Obtain parametric mapping of design parameters to reduced order coordinates
4) Predict high-dimensional data field at out-of-sample design points

Analysts must be able to select the appropriate procedures for each of these steps if accurate field surrogate models are to be generated.

For this study, training points and validation points are selected using maximin Latin Hypercube DoEs generated using JMP®, in which points are chosen a priori to evenly sample the parameter space. Fields are generated by evaluating each point in the DoE over a discretized spatial domain and storing the results in an array of vectors. For this study, only steady-state parametric problems are considered; extensions of these methods to transient problems could be the subject of future work.

The open-source Python framework Scikit Learn v. 0.22.1 is used to carry out DR [45]. Scikit Learn contains a number of modules capable of carrying out the DR techniques of interest and also contains a number of utility functions that can aid in developing more customized functionality. POD, ISOMAP, and LLE are included in this version of Scikit Learn and have been validated through a number of examples and studies. For POD, the only tuning parameter that is necessary to carry out DR is RIC. The "PCA" module in SciKit Learn computes the explained information of each mode based on the corresponding singular values, which can be used directly to select the modes that must be preserved to satisfy a desired RIC. For ISOMAP and LLE, the tuning parameters k, ϵ, and γ must be selected and do not have well established procedures for doing so a priori. To address this, an algorithm adapted from Reference [27] is implemented to optimize k, the pseudocode for which is presented in Algorithm 1. In this procedure, a metric, s, is computed that quantifies the quality of the dimensionality reduction and is minimized through optimization. Per recommendations in References [27] and [46], Kruskal's stress is minimized when selecting k for ISOMAP, and the variance of the ratio of Euclidean distance in the low-dimensional space to geodesic distance in the high-dimensional space is minimized when selecting k for LLE. These metrics are given in the following equations:

$$s_{kruskal}(k) = \sqrt{\frac{\sum(d_{ij}(\mathbf{W}^i, \mathbf{W}^j) - ||\mathbf{y}^i - \mathbf{y}^j||)^2}{\sum d_{ij}^2(\mathbf{W}^i, \mathbf{W}^j)}} \tag{13}$$

$$s_{var\ ratio}(k) = var\left(\frac{||\mathbf{y}^i - \mathbf{y}^j||_2}{d_{ij}(\mathbf{W}^i, \mathbf{W}^j)}\right) \tag{14}$$

Where d_{ij} is the shortest path distance along the k-nearest neighbors graph between two points (meant to approximate geodesic distance along the manifold) and $\mathbf{y}$ is the vector of coordinates of the corresponding points in the low-dimensional space. When performing the reconstructions for ISOMAP and LLE, The reconstruction weights must sum to one, as stated in Equation 11. For poorly chosen hyperparameters, it is possible for reconstruction weights to have large positive or negative magnitudes, which can result in spurious oscillations at grid points near a discontinuity and can result in non-physical predictions, such as negative densities. To mitigate this issue, an additional step is added to

the optimization in which each validation point is evaluated and the corresponding reconstruction weight coefficients are computed. If all weights are within the interval $[-0.25, 1.25]$, then the value of k is accepted. Otherwise, k is incrementally increased until the weight constraint is achieved or a maximum number of iterations is reached (for this study, the maximum number of iterations is chosen to be 10). The decision to incrementally increase k is based on the engineering experience that DR quality does not degrade significantly for modest increases in k starting from the optimal value, k^*, and increasing k can often have a smoothing effect on w since more points are available to perform the reconstruction. The procedure carried out in Algorithm 1 only optimizes k and assumes given values of ϵ and γ. To address this, reasonable ranges of ϵ and γ are chosen based on engineering experience and models are trained using each combination of the min and max values of each parameter. Based on previous experience, model quality is much more sensitive to k than it is to ϵ or γ, so the implementation of a robust procedure for selecting k is prioritized. A more efficient procedure that optimizes all relevant hyperparameters may be developed in future studies.

As mentioned previously, an unstructured interpolator is needed to map design parameters to the low-dimensional coordinate space. Radial Basis Function (RBF) interpolators were identified in the literature as plausible candidates since they meet this criterion. The RBF interpolator that is built in to Scipy v. 1.2 is used to generate interpolators in this research [47]. RBFs can be trained with a wide variety of possible kernel functions; for this study, the Thin-Plate Spline (TPS) and Gaussian kernels are considered. An unstructured linear interpolator is also considered for comparison.

Different back-mapping methods must be developed for each method. The "PCA" module in Scikit Learn contains methods for inverting POD for a given point in the low-dimensional coordinate space, so no additional development effort is necessary for POD implementation. For ISOMAP and LLE, the back-mapping procedure discussed in the previous section is implemented in Python 3.7.

To evaluate the quality of field predictions, it is necessary to derive an error metric quantifies the prediction error over the entire grid. To do so, an error metric defined in Reference [27] is used that aggregates the prediction error over each point in the field into a scalar quantity that can be directly compared across predictions using the following equation:

$$error(\mathbf{W}^{val}(\mathbf{p}^*), \mathbf{W}^*(\mathbf{p}^*)) = \frac{\sum_i |W_i^{val} - W_i^*|}{\sum_i |W_i^{val}|} \tag{15}$$

Where $\mathbf{p}^*$ is the set of design parameters at the validation point, $\mathbf{W}^{val}$ is the true field snapshot obtained by evaluating the FOM at $\mathbf{p}^*$, $\mathbf{W}^*$ is the predicted field at $\mathbf{p}^*$, and the subscript i corresponds to the i^{th} grid point in the field. This metric can be computed over the entire grid, or can be applied to specific regions of interest in the field, such as in the vicinity of a discontinuity. In this study, the "vicinity of the discontinuity" is obtained one of two ways. For analytical test problems in which step sizes are constant and known, the vicinity of the shock is defined as any grid point whose value does not correspond to one of the two "steady state" values observed on either side of the shock. Stated formally, $\mathcal{I} \coloneqq \{i \in \{1, ..n\} | W_i \neq f_{s1}, f_{s2}\}$, where $\mathcal{I}$ is the set of indices of grid points that are within the vicinity of the shock, W_i is the i^{th} component of the field vector $\mathbf{W}$, and f_{s1} and f_{s2} are the observed steady values of the function before and after the discontinuity respectively. For CFD problems in which the post-shock values are not constant and not known a priori, an engineering approximation is used to infer which points which points are in the vicinity of the shock by selecting gird points whose deviation between the predicted value and the true value exceed some threshold, τ. This approximation is similar to a procedure implemented in Reference [27]. For this study, $\tau \coloneqq 0.01(max_i \mathbf{W}^{val} - min_i \mathbf{W}^{val})$ is used to define the threshold. The set of points in the vicinity of the shock is then defined as follows: $\mathcal{I} \coloneqq \{i \in \{1, ..n\} | \, |W_i^{val} - W_i^*| > \tau\}$. Shock error is then computed by summing Equation 15 over the set $\mathcal{I}$, rather than summing over all grid points.

IV. Results

A. Test Problem: Simple Step Function

Perhaps the most challenging aspect of the application of ROMs to hypersonic flows is the representation and reconstruction of shocks. To isolate this issue, it is useful to apply ROM techniques to a simple discontinuous step function based on the following analytical definition:

$$f(x; p) = \begin{cases} 0 & 0 \leq x < l(p) \\ 1 & l(p) \leq x \leq 1 \end{cases} \tag{16}$$

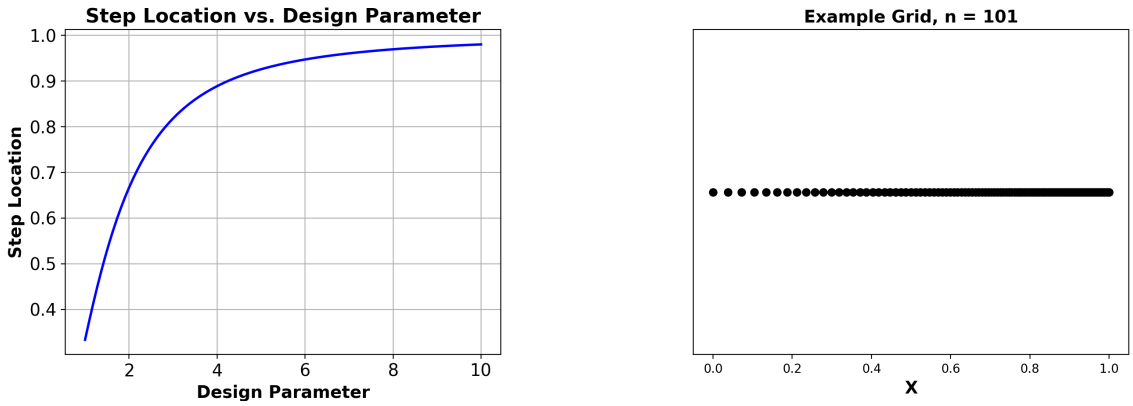
Algorithm 1: Optimize Number of Nearest Neighbors k (Adapted from Reference [27])

Data: Training cases stored in $m \times n$ array; Validation cases stored in $v \times n$ array; User-defined ϵ and γ; Metric function $s(\cdot)$

Result: Optimized value of k

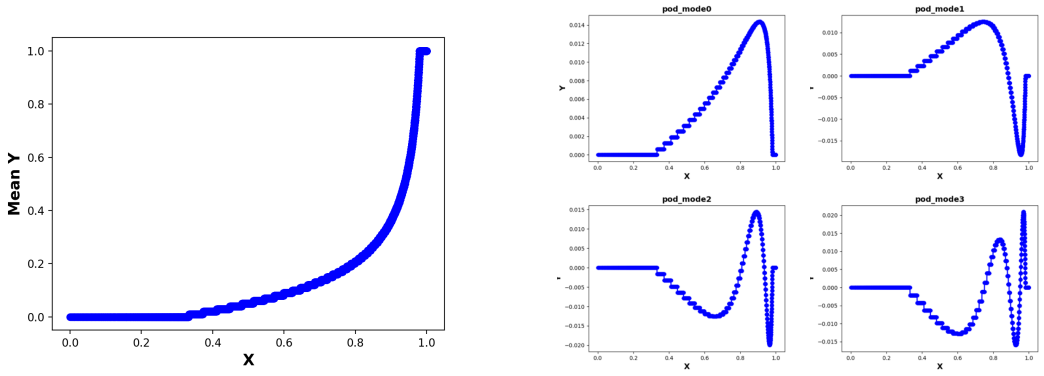
Compute optimal embedding;
$k_{list} = [(dim(\mathbf{p}) + 1) \ : \ m - 1]$; $s^* = k^* = \infty$;
while $|k_{list}| > 1$ **do**
    $k_{test} = \left[\lfloor \frac{1}{4}|k_{list}| \rfloor, \lfloor \frac{1}{2}|k_{list}| \rfloor, \lfloor \frac{3}{4}|k_{list}| \rfloor\right]$;
    $s_{test} = [0, 0, 0]$;
    for i *in* k_{test} **do**
        Apply DR at $k = k_{list}[i]$;
        $s_{test}[i] = s(k_{list}[i])$;
        if $s(k_{list}[i]) < s^*$ **then**
            $s^* = s(k_{list}[i])$;
            $k^* = k_{list}[i]$;
    if $min(s_{list}) == s_{test}[0]$ **then**
        delete $k_{list}\left[\lfloor \frac{1}{2}|k_{list}| \rfloor :\right]$;
    else if $min(s_{list}) == s_{test}[1]$ **then**
        delete $k_{list}\left[\lfloor \frac{3}{4}|k_{list}| \rfloor :\right]$;
        delete $k_{list}\left[0 : \lfloor \frac{1}{4}|k_{list}| \rfloor\right]$;
    else
        delete $k_{list}\left[0 : \lfloor \frac{1}{2}|k_{list}| \rfloor\right]$;

Check bounds on reconstruction weight coefficients $k_{best} = k^*$; $w_{min} = -0.25$; $w_{max} = 1.25$; $max\_iter = 10$;
for i *in* $range(v)$ **do**
    Compute weight coefficients for prediction i;
Obtain min/max weight coefficient for prediction i;
if $w_{min_i} > w_{min}$ *and* $w_{max_i} < w_{max}$ **then**
    return k^*
else
    $w^*_{min} = \infty$; $w^*_{max} = -\infty$; $iter = 0$; $k = k^*$;
    while $w_{min_i} < w_{min}$ *and* $w_{max_i} > w_{max}$ **do**
        k += 1;
        Apply DR using k;
        for i *in* $range(v)$ **do**
            Compute weight coefficients for prediction i;
            Obtain min/max weight coefficient for prediction i;
        if $w_{min_i} > w_{min}$ *and* $w_{max_i} < w_{max}$ **then**
            return k
        $iter$ += 1;
        if $iter > max\_iter$ **then**
            Warn user that optimal k could not be found to meet criteria;
            return k^*



(a) Step Location vs. Design Parameter **(b) Low-Dimensional Representation**

Fig. 2 Setup of 1-D Step Function Test Problem



(a) Sample Mean **(b) Four Most Significant Modes**

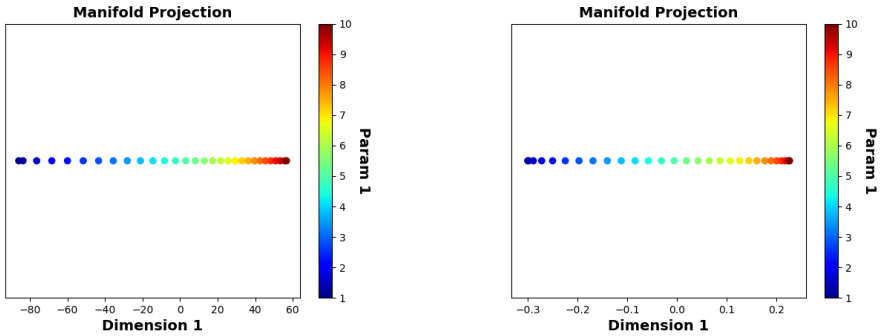
Fig. 3 POD Mean and Four Most Significant Modes

$$l(p) = \frac{p^2}{p^2 + 2} \tag{17}$$

Where p is a user-defined input parameter that controls the location of the step, $l(p)$, on the interval $x \in [0, 1]$. By varying the parameter p, snapshots can be generated with different step locations based on Equation 17. The relation between p and the shock location $l(p)$ is shown in Figure 2a and an example grid over which the snapshots are generated is shown in Figure 2b. This problem is useful for illustrative purposes because it eliminates much of the complexity that exists in real hypersonic applications so focus can be placed specifically on the ability to capture discontinuities.

Snapshots are generated by sampling m evenly-spaced values of $p \in [1, 10]$. The grid is composed of n points along the interval $x \in [0, 1]$; the grid points are chosen to be a higher resolution in x than the snapshots have in p to avoid samples where different values of p result in identical snapshots. The grid is generated using logarithmic spacing to provide high grid resolution near $x = 1$, where steps are likely to cluster if p is sampled evenly. Snapshots are generated by storing the values of f at each grid point in x in a vector to form the data set $F = \{\mathbf{f}_1, ..., \mathbf{f}_m\}$, where $\mathbf{f}_i \in \mathbb{R}^n$ and the subscript i corresponds to the value p_i used to generate the i^{th} snapshot.

POD is then applied to the data set F to obtain a low-dimensional linear projection. Since $m < n$, POD projects the data into a linear subspace spanned by $m - 1$ orthogonal modes, though this number can be further reduced by applying an $RIC < 1$ based on Equation 5. To illustrate this, Figure 3 shows the sample mean and the four most significant modes that are produced by applying POD to the problem for the case where $m = 100$ and $n = 1000$.



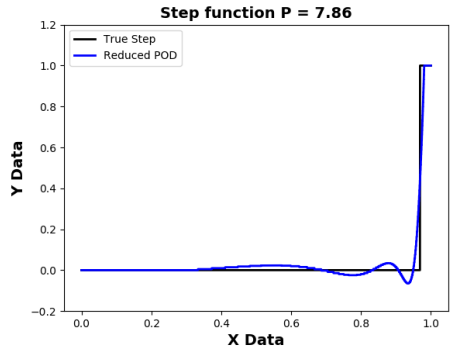
(a) ISOMAP Embedding **(b) LLE Embedding**

Fig. 4 Low-Dimensional Representations Obtained Using Manifold Learning Methods

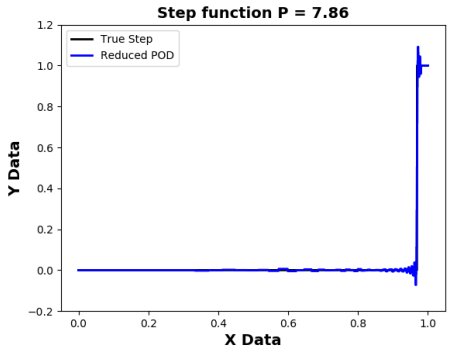
To assess the impact that POD model reduction has on model accuracy, $RIC = 0.9, 0.99, 0.999$, and 1.0 are considered. POD models are trained using each value of RIC to reduce the number of modes that are preserved in the projection. For each model, the preserved modes are identified and interpolate models are fit to map values of p to the coefficients a_j that correspond to each mode $\phi_j(x)$. These coefficients are then used to reconstruct the high-dimensional data at a prediction point, p^* by linearly combining modes according to Equation 6.

After the POD-bsed field surrogate models are generated, models are developed using ISOMAP and LLE. When applying these NLDR algorithms, it is appropriate to assume that the dimension of the underlying manifold is one, since it is known a priori that only one parameter is being varied to generate the snapshots [26]. Using the optimization procedure discussed in the previous section, it is found that $k = 2$ is the optimal number of nearest neighbors for both ISOMAP and LLE. Due to the simplicity of this problem, predictions for each model were found to be largely insensitive to ϵ and γ, so default values of $\epsilon = 0.01$ and $\gamma = 2$ are used. The embeddings obtained by each algorithm are show in Figure 4. Both algorithms are able to obtain an embedding that accurately reflects the behavior in the original data set. Not only are embedded points sorted by design parameter, both embeddings also exhibit a clustering of points near $p = 10$, which is due to the clustering of steps across snapshots (illustrated in Figure 2a). When comparing this representation to the modes recovered by POD in Figure 3, it is shown that both ISOMAP and LLE are able to obtain a low-dimensional representation that is more reflective of the behavior observed in the training set, which could impact the relative quality of models trained using each method. Using these embeddings, the back-mapping procedure discussed in the previous section is then used to reconstruct the high dimensional data at predicted values.

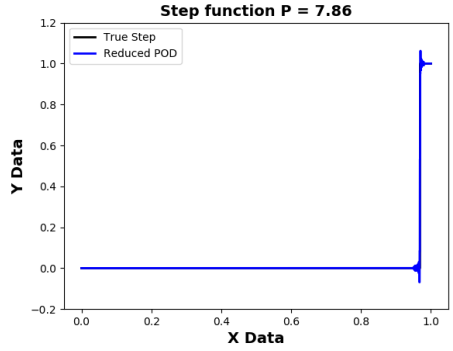
Once all surrogate models are generated, 20 validation points are evaluated using each model. These validation points are chosen to be evenly spaced along the interval $p \in [1.2, 9.8]$. The accuracy of the predictions made by each model are compared to assess the accuracy of each method. For illustration, predictions made by each model are shown in Figure 5 at $p = 7.86$ for $m = 100$ and $n = 1000$. It is observed that the reduced POD models (Figures 5a-5c) approximately capture the location of the step, but the function values near the step exhibit spurious oscillations which are highly undesirable. Even for $RIC = 0.999$, oscillations in the vicinity of the step can exhibit fluctuations larger than $\pm 10\%$. Meanwhile, models trained using POD that preserve all available modes ($RIC = 1.0$), ISOMAP and LLE each accurately predict the step location with no spurious oscillations (Figure 5 only shows one example, but this behavior is observed over all 20 validation points). From this, it is possible to conclude that full POD, ISOMAP, and LLE are all viable candidate DR methods for discontinuous fields. Reduced POD ($RIC < 1.0$) is not considered for the remaining test applications.



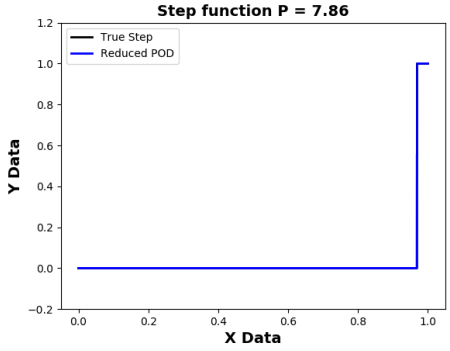
(a) POD, $RIC \geq .9$



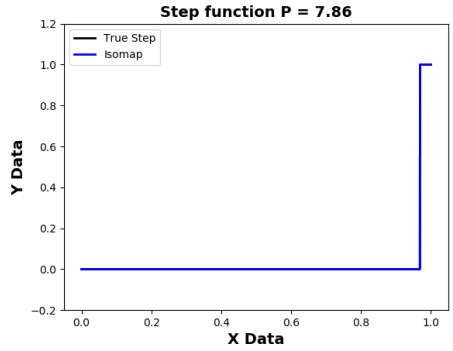
(b) POD, $RIC \geq .99$



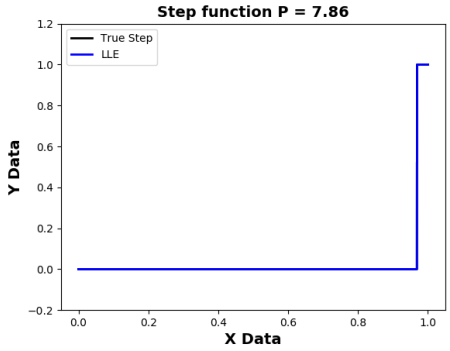
(c) POD, $RIC \geq .999$



(d) POD, $RIC = 1.000$

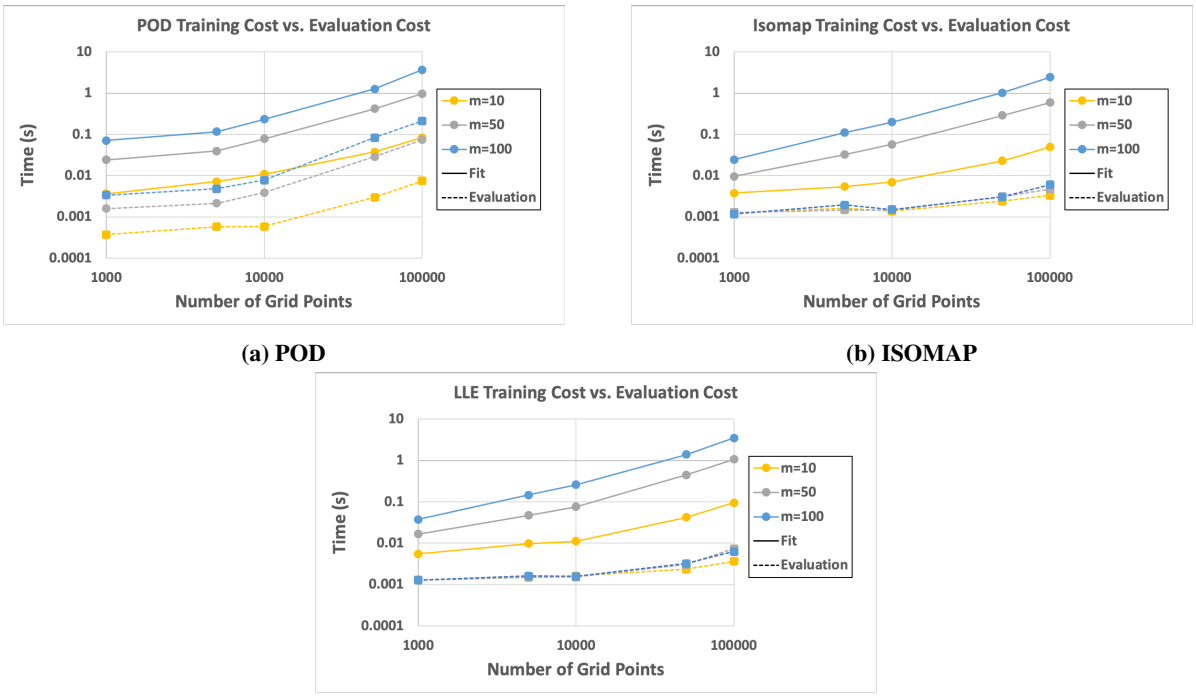


(e) ISOMAP



(f) LLE

Fig. 5 Example Step Reconstructions at $p = 7.86$**, Models Trained with Sample Size** $m = 100$



(a) POD **(b) ISOMAP**

(c) LLE

Fig. 6 Model Fit and Evaluation Costs vs. Model Type

While full POD, ISOMAP, and LLE are all found to produce surrogate models of comparable accuracy, it is important to note that the POD-based model must preserve $m-1$ dimensions while ISOMAP- and LLE-based models must only preserve one dimension. This has significant ramifications for evaluation cost, as is shown in Figure 6. To generate these plots, models are fit repeatedly using each method for a range of m and n. Training cost and average evaluation cost are measured in CPU-time over these ranges to assess trends and sensitivities. From these plots, it is observed that the evaluation cost of the POD model increases rapidly as both number of training snapshots, m, and number of grid points, n, increase. This is a direct consequence of the fact that full POD must preserve $m-1$ DoFs to adequately resolve the step. By evaluating an $m-1$ dimension interpolate model, and combining $m-1$ modes of dimension n when generating predictions, the evaluation cost of the full POD-based model increases rapidly as the training data set increases in size. Meanwhile, the evaluation costs of ISOMAP- and LLE-based models are both shown to be insensitive to m, and grow much more modestly as n increases. This is because ISOMAP and LLE both produce models with only one DoF, and only require $k=2$ snapshots to be combined when generating predictions. Low evaluation cost is critical to the overall motivation of implementing these models into coupled analyses. The differences in scaling between linear methods and the tested nonlinear methods is significant and could be used to justify the application of nonlinear ROMs in future studies.

B. Test Problem: Two-Dimensional Step Function

The previous test problem is intended to isolate and compare the ability of POD, ISOMAP, and LLE to represent variable discontinuities over a grid in the simplest case. Next, a test problem is derived to increase the complexity of the problem to include a few more distinct features of hypersonic applications that are of interest. The test problem considered here is a two-dimensional step function along a wedge with variable angle, as shown in Figure 7. The value of the scalar function is given by the following equations:

$$f(x,y;a,\beta,\delta) = \begin{cases} a^2 & y \le s(x;\beta,\delta) \\ 0 & y > s(x;\beta,\delta) \end{cases} \tag{18}$$

$$s(x;\beta,\delta) = tan(\beta+\delta) * x \tag{19}$$

Where a controls the size of the step, β is the angle of the step relative to the wedge, and δ is the angle of the wedge. The geometry of this problem is intended to resemble a simplified version of two-dimensional supersonic flow over a wedge with a "shock", but simplified using analytical functions to eliminate complexities of physical flow solutions and allow for more direct control over features of interest.

The two key features of interest that this test problem adds are the parametric nonlinear variation of step size and the variation of grid geometry. The spatial domain is discretized using a 200×200 structured grid, in which grid points are arranged in an ordered manner such that each grid point can be referenced using i, j indices, as illustrated in Figure 8. The grid is fixed to the wedge, so varying the wedge angle, δ, causes the spatial location of grid points to change. To train the model, a Latin Hypercube DoE of 200 samples is evaluated using the parameter ranges listed in Table 3. An additional Latin Hypercube DoE of 50 samples are evaluated and used as validation samples to evaluate the predictive capabilities of the model.

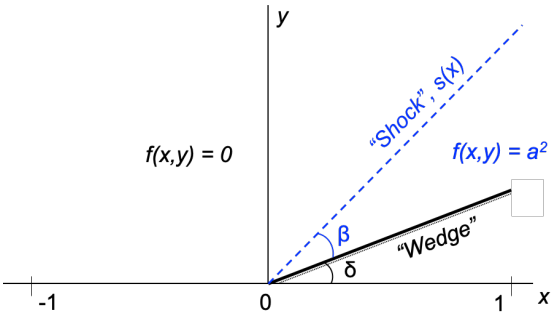


Fig. 7 Illustration of Two-Dimensional Test Problem

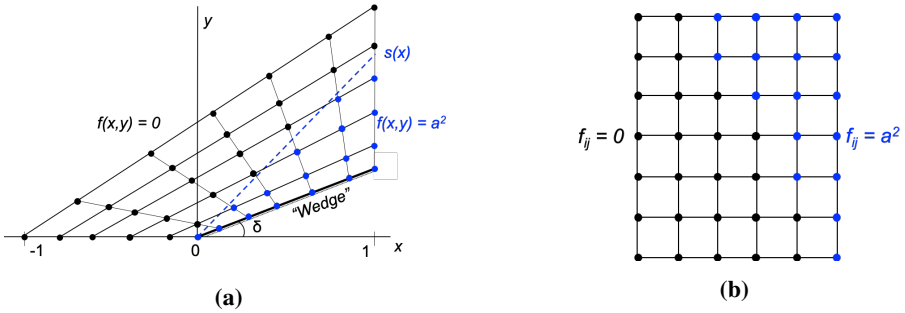


Fig. 8 Notional Illustration of the (a) Grid Fit to a Sample Two-Dimensional Geometry and the (b) Transformation to a Common Structured Format

After the training samples are generated, field surrogate models are trained. POD, ISOMAP, and LLE are considered as candidate methods for dimensionality reduction. Additionally, linear interpolation, Thin-Plate Spline (TPS) RBF, and Gaussian RBF methods are considered as candidate methods for the interpolation step. For this problem, POD models are only trained using $RIC = 1.0$ to maximize model accuracy. For ISOMAP and LLE, the hyperparameters k, ϵ, and γ are optimized using the procedure discussed in Section III. Due to the categorical nature of these variables, a full-factorial combination of these methods are used to train models, resulting in 3 unique POD-based models (3

Parameter	Min	Max
a	2	8
β	5°	20°
δ	5°	15°

Table 3 Input Parameter Ranges for 2-D Analytical Step Function

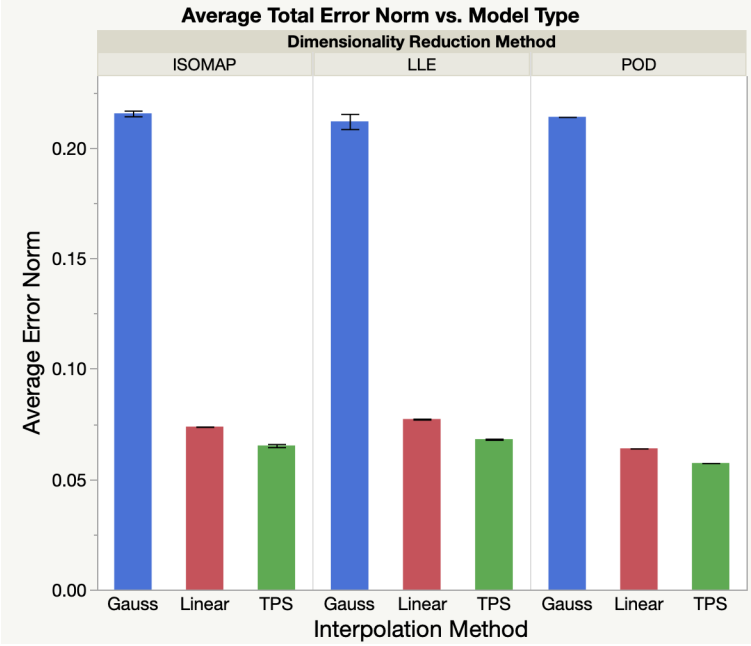


Fig. 9 Average Norm of Error over Full Grid vs. Model Type

interpolation methods), 12 unique ISOMAP-based models (3 interpolation methods $\times$ 2 ϵ values $\times$ 2 γ values) and 12 unique LLE-based models (same combinations as ISOMAP).

After each model is trained, the 50 validation points are evaluated using each model and error metrics are computed. As discussed in the previous section, the norm of the error can be computed to reduce error over the field to a scalar quantity which can then be compared across models. For this problem, it useful to look at the error norm both over the entire field (referred to as "total error") and in the vicinity of the step (referred to as "shock error"). To isolate the grid points near the step, the "steady" values of the function are obtained by sampling points in the field that are far from the discontinuity and are not affected by near-shock errors or oscillations in the prediction. Grid points with these "steady" values are omitted when computing the shock error. The error norms over all 50 validation points are averaged and used to compare each model. The error norms over the full grid and in the vicinity of the shock for each model type are shown graphically in Figure 9 and 10 respectively. In these plots, the values of the main bars correspond the average value of the average error norm for models trained using each combination of dimensionality reduction and interpolation technique. The interval for the bars in the ISOMAP and LLE groups correspond to the full range observed due to the variations of ϵ and γ.

Parameter	Min	Max
k	4	199
ϵ	0.01	0.1
γ	2	4

Table 4 NLDR Hyperparameter Ranges

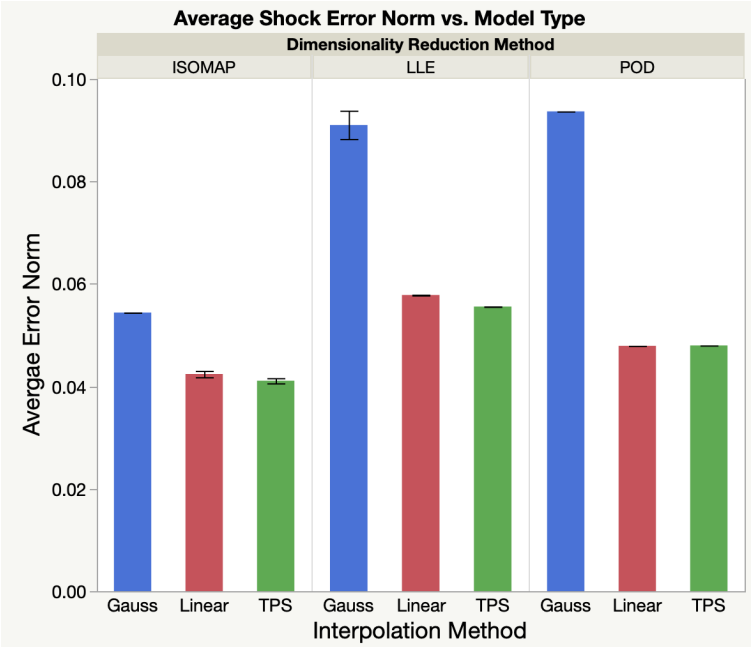


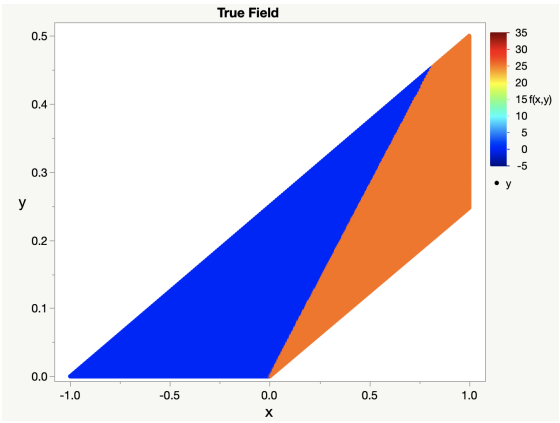
Fig. 10 Average Norm of Error near Shock vs. Model Type

Method	Interpolator	k	ϵ	γ
POD	RBF (TPS)	–	–	–
ISOMAP	RBF (TPS)	20	0.1	4
LLE	RBF (TPS)	194	0.1	2

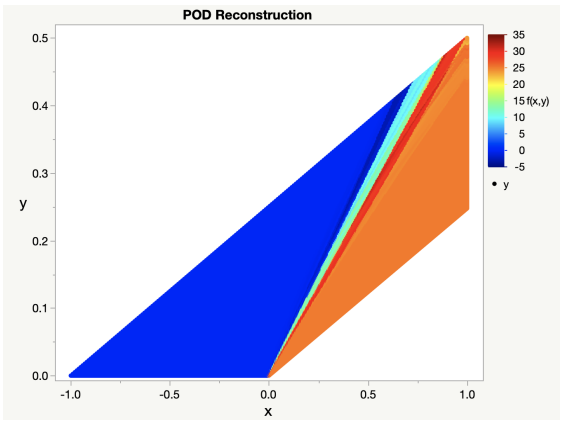
Table 5 Best Model Settings for each DR Technique

Model (best)	Avg. Undershoot	Avg. Overshoot	Avg. Error Norm Over Grid	Avg. Error Norm Near Shock	Avg. Step Size Error
POD	-17.6%	40.2%	0.0573	0.0479	1.55%
ISOMAP	-10.7%	11.9%	0.0645	0.0406	2.62%
LLE	-1.22%	4.83%	0.0679	0.0579	1.92%

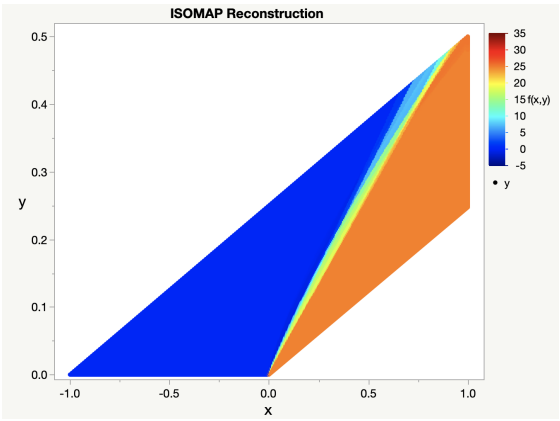
Table 6 Error Metrics for 2-D Step Predictions



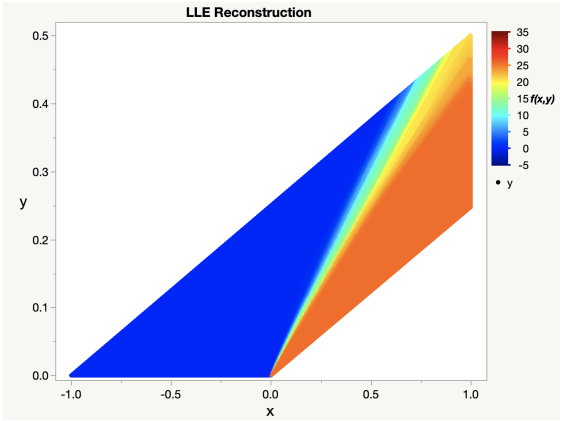
(a) True Field



(b) POD



(c) ISOMAP



(d) LLE

Fig. 11 Example 2D Step Reconstructions at $a = 5.06$, $\beta = 10.5°$, and $\delta = 13.9°$

The settings that produce the best models for each DR technique are shown in Table 5. The error metrics that correspond to these models are shown in Table 6. An example of predictions for each of these models is shown in Figure 11. The cost of training and evaluating these models is reported in Table 7. From the data, several important observations can be made. First, it is shown that POD outperforms both ISOMAP and LLE when predicting the size of the step. This is observed by comparing the average error in the "steady" prediction of the step size and is also reflected in the computation of the error norm over the full grid. Next, it is observed that ISOMAP outperforms both POD and LLE when predicting behavior in the vicinity of the step. This is because ISOMAP exhibits less extreme oscillations near the step, and the errors that ISOMAP predictions exhibit occur in a narrower region of space. This is observed qualitatively in Figure 11 and is quantified by comparing the overshoot/undershoot of the model. The error norm in the vicinity of the step is lowest using ISOMAP; however, the error norm over the full grid is lowest for POD since ISOMAP accumulates larger error in the "steady" region of the field due to a larger discrepancy in the prediction of the step size. Finally, LLE under-performs POD and ISOMAP in all metrics. This is likely because LLE models predict

Method	Time per DR Call [CPU-s]	Total Training Time [CPU-s]	Time per Prediction [CPU-s]
POD	2.198	2.630	0.097
ISOMAP	2.540	97.65	0.022
LLE	124.16	2879.98	0.228

Table 7 Cost Data for Best 2-D Step Function Models

Parameter	Min	Max
M	4	6
α	-2°	4°

Table 8 Input Parameter Ranges for Fixed Geometry Airfoil

very smooth steps that exhibit variation over a large region of space. While LLE models exhibited the best overshoot and undershoot and outperformed ISOMAP in the prediction of the step size, the prediction of behavior near the step is too smooth and highly undesirable when trying to resolve discontinuous features. The smoothing effect observed in the LLE models is likely caused by the high value of k found by the optimizer and could be specific to the nature of this problem. For this problem, the cost of a single DR call is comparable between ISOMAP and POD. However, the total training cost of the ISOMAP-based model is mach larger than the cost of training the POD-based model because DR must be performed repeatedly for the ISOMAP-based model when optimizing k while the POD must only be performed once. The evaluation cost of the ISOMAP-based model is significantly lower than the evaluation cost of the POD-based model. Based on these values, the ISOMAP-based model would break even with the POD-based model after approximately 1267 evaluations. This number could be reduced if a more efficient k-optimization scheme is developed, which will be the subject of future investigation. The evaluation training cost of the LLE model is substantially higher than ISOMAP or POD because of the high value of k found by the optimizer. LLE is impractical based on these cost and performance data, but these short-comings may be specific to the nature of this problem.

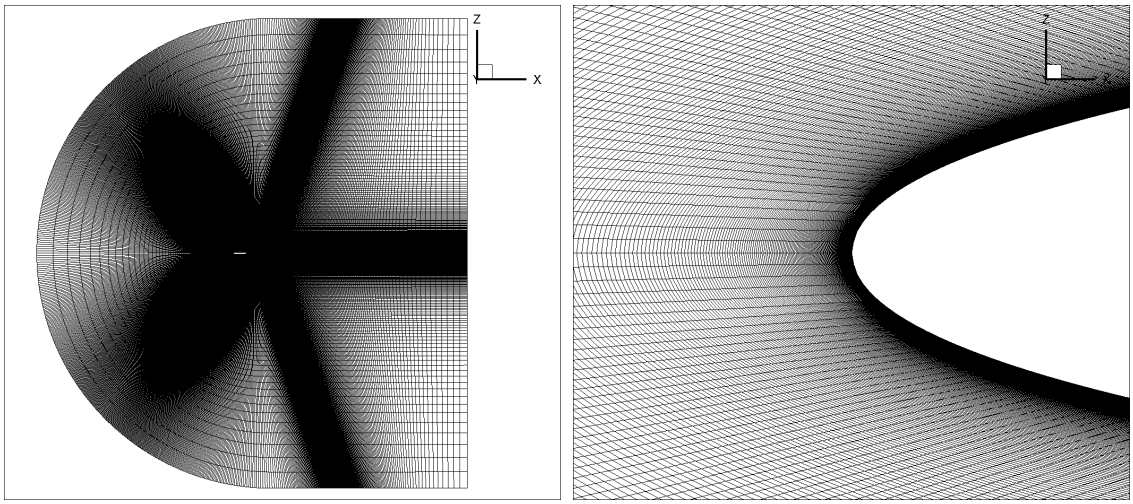
C. Two-dimensional Hypersonic Flow over Airfoil

The final test problem considered in this study is the application of field surrogate modeling to the prediction of high-speed aerodynamic fields. In this problem, the aerodynamic field is computed over a symmetric NACA 0006 airfoil. To generate solutions, NASA's Fully Unstructured Navier-Stokes 3D (FUN3D) v. 13.4 is used. FUN3D was developed at the NASA Langley Research Center and is the current state-of-the-art flow analysis support tool for NASA programs [48]. FUN3D is an unstructured, node-based Euler and RANS solver that can be applied to two- or three-dimensional flows for either transient or steady-state conditions. For this study, solutions are computed using steady-state, laminar, compressible, perfect gas physics. The solver is executed with a van Leer flux splitting method with a heuristic van Albada flux limiter for the inviscid flux calculations. The solver is run for 4000 total iterations; the first 50 iterations are carried out using first order approximations and the last 200 iterations are carried out with a frozen limiter. Solutions are considered converged if the residual for each of the governing equations decreases by three orders of magnitude. Altitude is assumed to be $30,000\ m$; reference quantities and scaling parameters are computed based on the ambient properties at this altitude according to the ARDC standard atmospheric model [49].

Grid generation is carried out using Pointwise® v. 18.2 [50]. Pointwise allows for the generation of both structured and unstructured grids and contains a scripting interface that allows gird generation to be partially automated. In this study, snapshots are generated using structured grids to ensure that the number of grid points and their relative locations and connectivity are preserved across a set of snapshots. This ensures that each grid point has the same behavior relative to each other grid point across each snapshot in the data set, which allows the DR algorithm to recover meaningful patterns in the data set [26, 35]. It is important to note that methods have been developed to deform unstructured grids so variable geometries with unstructured grids can have these same properties [51, 52]. However, these methods present significant challenges as they are often limited in the ranges of shape variation that can be realized before the deformed grid degenerates, especially when fine, high aspect ratio cells are used to resolve boundary layers [35]. Though deformable unstructured grids could potentially be feasible for some applications, they are not used in this study for logistical purposes.

The domain of the field around the airfoil is discretized using a 1197×400 structured grid, as is shown in Figure 12. To generate the training data, a Latin Hypercube DoE of 100 samples is generated by varying flow Mach number (M) and angle of attack (α) using the variable ranges listed in Table 8. An additional 50 validation cases are generated using a Latin Hypercube DoE with the same ranges. The same model types, combinations, and procedures are applied to this test problem as were applied to the 2-D analytical step function in the previous section.

After each model is trained, the 50 validation points are evaluated using each model and error metrics are computed. The ranges of error norms over the entire grid and in the vicinity of the shock are presented in Figures 13 and 14 respectively. The settings that produce the best models for each DR technique are listed in Table 9. Using these best



(a) Full Grid **(b) Zoomed in on Leading Edge**

Fig. 12 Structured Airfoil Grid for NACA 0006 Airfoil Test

Method	Interpolator	k	ϵ	γ
POD	RBF (TPS)	–	–	–
ISOMAP	RBF(TPS)	15	0.01	4
LLE	RBF (TPS)	24	0.01	4

Table 9 Best Model Settings for each DR Technique

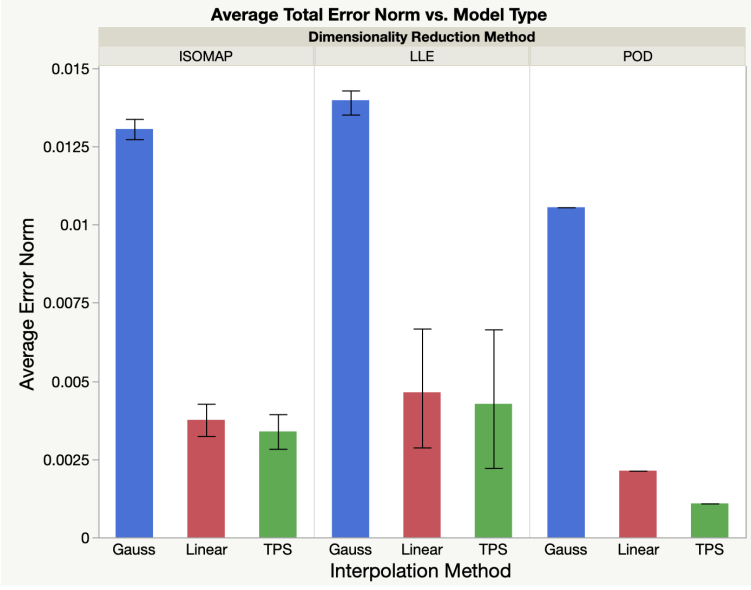


Fig. 13 Average Norm of Error over Full Grid vs. Model Type

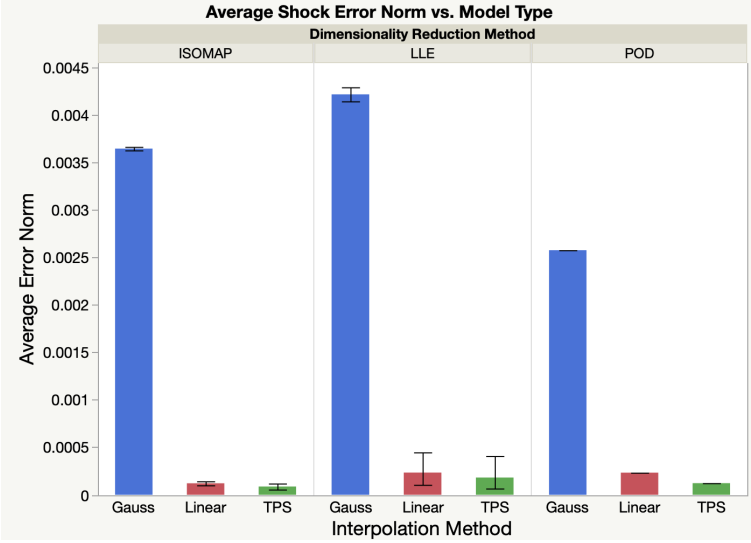


Fig. 14 Average Norm of Error near Shock vs. Model Type

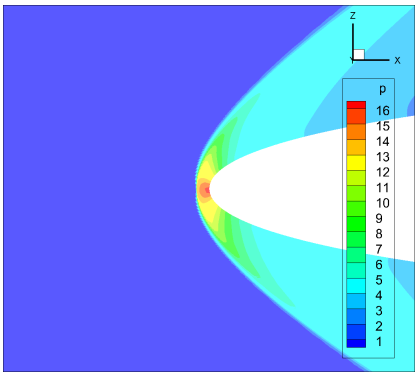
Method	Time per DR Call [CPU-s]	Total Training Time [CPU-s]	Time per Prediction [CPU-s]
POD	22.10	23.55	1.125
ISOMAP	17.71	360.5	0.0930
LLE	36.46	1513.4	0.1457

Table 10 Cost Data for Airfoil Models

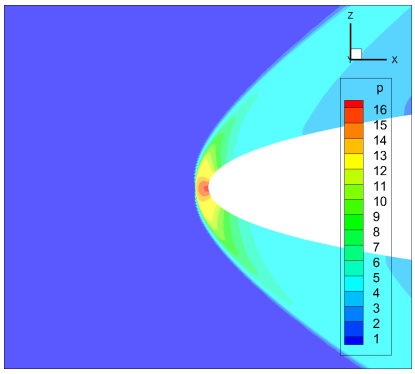
model settings, an example prediction and the associated absolute error calculations for the pressure field at $M = 4.2$ and $\alpha = 3.02°$ are depicted in Figure 15. The cost of training and evaluating these models is reported in Table 10. The observations that can be drawn from these figures are similar to those obtained in the 2-D analytical step function. Again, POD-based models exhibit lower error than ISOMAP and LLE over the entire grid while ISOMAP- and LLE-based models outperform POD when making predictions in the vicinity of the shock. The optimal value of k found for the LLE-based model is much lower for this problem than in the 2-D analytical step function, which enabled a more precise reconstruction of the discontinuity. For all DR methods, an RBF interpolator with a TPS kernel is found to produce predictions with the lowest error while the RBF interpolator with a Gaussian kernel produces predictions with a significantly higher error. LLE-based models are found to be much more sensitive to variations in ϵ and γ than the ISOMAP-based models. Similar to the 2-D step function, the training cost of the POD-based model is found to be lower than the ISOMAP- and LLE-based models, but the ISOMAP- and LLE-based models have significantly lower evaluation costs. Using the values in Table 10, it is estimated that the total cost of the ISOMAP-based model would break even with the POD-based model after approximately 327 evaluations, while the total cost of the LLE-based model would break even with the POD-based model after approximately 1522 evaluations. Compared to the 2-D step function, the lower value of k that is obtained for the LLE-based model results in a much more reasonable cost, meaning that the practical application of these ROMs to coupled analyses is highly dependent on the choice of k.

V. Conclusion

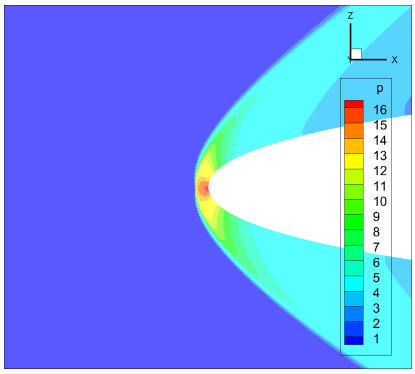
The goal of this paper is to make meaningful contributions to the state-of-the-art in the field of hypersonics systems engineering and MDAO. Reduced order models have been identified as a key enabler to allow hypersonic system designers to represent high-fidelity field data in coupled analyses at a manageable evaluation cost. This study describes the implementation of a ROM-based design methodology developed by implementing multiple techniques from the open literature. The experiments of this study demonstrate the ability of this methodology to generate field surrogate models at the disciplinary level. Future studies will focus on expanding the suite of methods available for each step of



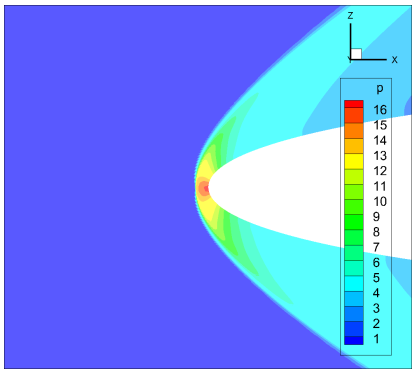
(a) True Solution



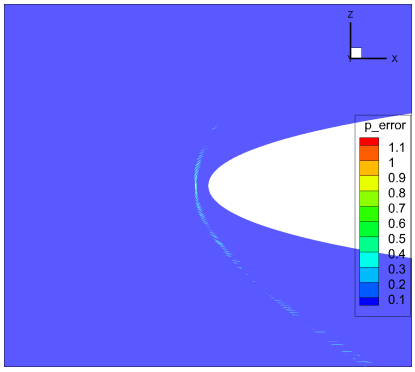
(b) POD Prediction



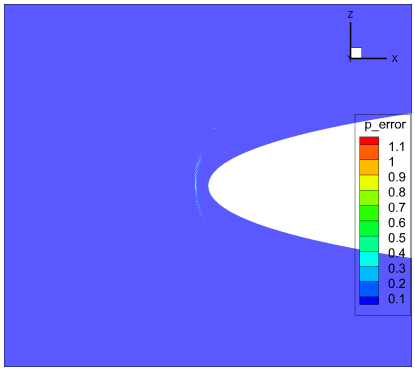
(c) ISOMAP Prediction



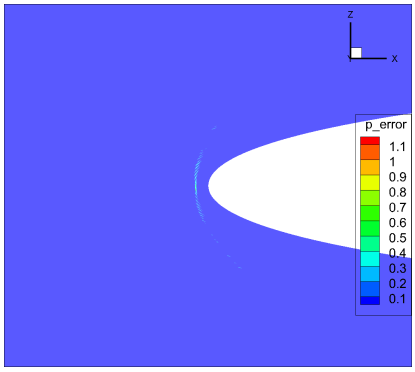
(d) LLE Prediction



(e) POD Error



(f) ISOMAP Error



(g) LLE Error

Fig. 15 Predictions of Pressure near Airfoil Leading Edge and Absolute Prediction Errors for DR Method

the process, characterizing these methods to enable analysts to select the proper methods for a given application, and on the implementation of these methods into coupled analysis to demonstrate their capabilities and benefits when used to make system-level predictions.

Another contribution of this study is the direct comparison of linear DR (namely POD) with nonlinear DR (namely ISOMAP and LLE) in terms of their ability to generate accurate and cost effective field surrogate models. In the simple step function test problem, it is observed that reduced POD is unable to resolve discontinuous features without spurious oscillations but was able to accurately predict step location when all DoFs are preserved. The 1-D step function also illustrates the ability of NLDR algorithms to recover clustering in the low-dimensional manifold that results from common behavior in the response fields across each sample. This could allow NLDR models to benefit from adaptive sampling, which has been shown to be promising in the literature will be explored in future work [27]. In the 2-D step function, it is observed that ISOMAP outperforms POD when resolving discontinuous features in the field while POD outperforms both ISOMAP and LLE in predicting the steady-state response behind the step. This discrepancy may be addressed by applying different DR and/or back-mapping techniques and will be the subject of further investigation. Similar behavior is observed in a 2-D CFD test case, in which both ISOMAP and LLE are found to outperform POD when making predictions of field quantities in the vicinity of a strong bow shock while POD outperformed ISOMAP and LLE when assessing accuracy over the grid as a whole. Future studies will look into increasing the complexity of the CFD applications, such as allowing for shape variations and assessing problems with a higher-dimensional design space. Finally, ISOMAP and LLE are shown to have the potential to create models with a much lower evaluation cost than those trained using POD, but the need for multiple DR calls during hyperparameter optimization results in ISOMAP- and LLE-based models having a significantly higher training cost. The total offline cost, and thus the break-even point of these models, is highly sensitive to k. Methods for optimizing the value of k more efficiently will be investigated in future studies. However, due to the lower evaluation cost, the nonlinear methods tested in this study could be more desirable candidates for training models that will eventually be implemented in coupled MDAO applications.

References

[1] Raymer, D. P., *Aircraft Design: A Conceptual Approach*, American Institute of Aeronautics and Astronautics, 1992.

[2] Mavris, et al., "A Stochastic Approach to Multi-disciplinary Aircraft Analysis and Design," *36th Aerospace Sciences Meeting and Exhibit*, 1998.

[3] Maicke, B. A., and Majdalani, J., "Evaluation of CFD Codes for Hypersonic Flow Modeling," *46th AIAA Joint Propulsion Conference and Exhibition*, 2010.

[4] Bowcutt, K. G., "A Perspective on the Future of Aerospace Vehicle Design," *AIAA International Space Planes and Hypersonic Systems and Technologies*, 2003.

[5] Bowcutt, K. G., et al., "Advancements in Multidisciplinary Design Optimization Applied to Hypersonic Vehicles to Achieve Closure," *15th AIAA International Space Planes and Hypersonic Systems and Technologies Conference*, 2008.

[6] Notz, W., "Introductory Overview Lecture on Computer Experiments - Design," , Aug. 2006.

[7] Ripepi, M., et al., "Reduced-Order Models for Aerodynamic Applications, Loads and MDO," *CEAS Aeronautical Journal*, Vol. 9, No. 1, 2018, pp. 171–193.

[8] Papageorgiou, A., et al., "Multidisciplinary Design Optimization of Aerial Vehicle: A Review of Recent Advancements," *International Journal of Aerospace Engineering*, Vol. 2018, 2018.

[9] Anderson Jr., J. D., *Hypersonic and High Temperature Gas Dynamics*, 2nd ed., American Institute of Aeronautics and Astronautics, 2006.

[10] Chavez, F. R., and Schmidt, D. K., "Analytical Aeropropulsive/Aeroelastic Hypersonic Vehicle Model with Dynamic Analysis," *Journal of Guidance, Control, and Dynamics*, Vol. 17, No. 6, 1994.

[11] Schwartz, H. D., et al., "Comparison of Three Aerodynamic Analysis Software Packages Against the Army Nave Finner Projectile to Determine Fidelity Level," *AIAA SciTech Forum*, 2018.

[12] Vincenti, W. G., and Kruger, C. H., Jr, *Introduction to Physical Gas Dynamics*, Krieger Publishing Company, 1965.

[13] Hermann, R., "Hypersonic Non-Equilibirum Flow and its Thermodynamic Relations," *UARI Research Report No. 30*, 1965.

[14] Gazaix, A., et al., "Investigation of Multi-Disciplinary Optimisation for Aircraft Preliminary Design," *SAE International*, 2011.

[15] Baker, C. A., et al., "Study of a Global Design Space Exploration Method for Aerospace Vehicles," *8th Symposium on Multidisciplinary Analysis and Optimization*, 2000.

[16] Franceschini, G., and Macchietto, S., "Model-based design of experiments for parameter precision: State of the art," *Chemical Engineering Science*, Vol. 63, No. 19, 2008, pp. 4846–4872.

[17] Liefsson, L., et al., "Aerodynamic Shape Optimization by Space Mapping," *Surrogate-Based Modeling and Optimization*, edited by S. Koziel and L. Leifsson, Springer, 2013, pp. 213–245.

[18] Friedmann, P. P., and Powell, K. G., "Aeroelasticity, Aerothermoelasticity, and Aeroelastic Scaling of Hypersonic Vehicles," *University of Michigan Department of Aerospace*, 2004.

[19] Kinney, D. J., et al., "Predicted Convective and Radiative Aerothermodynamic Environments for Various Reentry Vehicles Using CBAERO," *44th AIAA Aerospace Sciences Meeting and Exhibit*, 2006.

[20] Ferguson, F., et al., "A Coupled Aerodynamic and Propulsive Performance Analysis of a Generic Hypersonic Vehicle," *AIAA Propulsion and Energy Forum*, 2015.

[21] Myers, R. H., Montgomery, D. C., and Anderson-Cook, C. M., *Response Surface Methodology*, 3rd ed., John Wiley & Sons, Inc., 2009.

[22] Izenman, A. J., "Introduction to Manifold Learning," *Wiley Interdisciplinary Reviews: Computational Statistics*, 2012.

[23] Bui-Thanh, T., Damodaran, M., and Wilcox, K., "Proper Orthogonal Decomposition Extensions for Parametric Applications in Compressible Aerodynamics," *21st Applied Aerodynamics Conference*, 2003.

[24] Falkiewicz, N. J., and Cesnik, C. E. S., "Proper Orthogonal Decomposition for Reduced-Order Thermal Solution in Hypersonic Aerothermoelastic Simulations," *AIAA Journal*, Vol. 49, No. 5, 2011.

[25] Iuliano, E., and Quagliarella, D., "Proper Orthogonal Decomposition, Surrogate Modelling and Evolutionary Optimization in Aerodynamic Design," *Computers and Fluids*, Vol. 84, 2013, pp. 327–350.

[26] Franz, T., et al., "Interpolation-Based Reduced Order Modeling for Steady Transonic Flows via Manifold Learning," *International Journal of Computational Fluid Dynamics*, 2014.

[27] Franz, T., "Reduced-Order Modeling for Steady Transonic Flows via Manifold Learning," Ph.D. thesis, Deutches Zentrum fur Luft- und Raumfahrt, 2016.

[28] Lucia, D. J., and Beran, P. S., "Projection methods for reduced order models of compressible flows," *Journal of Computational Physics*, Vol. 188, No. 1, 2003, pp. 252–280.

[29] Schlegel, M., and Noack, B. R., "On long-term boundedness of Galerkin models," *Journal of Fluid Mechanics*, Vol. 765, 2015, pp. 325–352.

[30] Chen, H., "Blackbox Stencil Interpolation Method for Model Reduction," Master's thesis, Massachusetts Institute of Technology, 2012.

[31] Xiao, D., et al., "A Parameterized Non-Intrusive Reduced Order Model and Error Analysis for General Time-Dependent Nonlinear Partial Differential Equations and its Applications," *Computer Methods in Applied Mechanics and Engineering*, Vol. 317, 2017, pp. 868–889.

[32] Tropea, C., Yarin, A. L., and Foss, J. F., *Springer Handbook of Experimental Fluid Mechanics*, Springer-Verlag Berlin Heidelberg, 2007.

[33] Kerschen, G., et al., "The Method of Proper Orthogonal Decomposition for Dynamical Characterization and Order Reduction of Mechanical Systems: An Overview," *Nonlinear Dynamics*, Vol. 41, No. 1-3, 2005, pp. 147–169.

[34] Sirovich, L., "Turbulence and the Dynamics of Coherent Structures: Part I: Coherent Structures," *Quarterly of Applied Mathematics*, Vol. XLV, No. 3, 1987, pp. 561–571.

[35] Bozeman Jr., M. D., "A Reduced Order Modeling Methodology for the Multidisciplinary Design and Analysis of Boundary Layer Ingestion Configurations," Ph.D. thesis, Georgia Institute of Technology, 2019.

[36] Crowell, A. R., and McNamara, J. J., "Model Reduction of Computational Aerothermodynamics for Hypersonic Aerothermoelasticity," *AIAA Journal*, Vol. 50, No. 1, 2012, pp. 74–84.

[37] Crowell, A. R., et al., "A Reduced Order Aerothermodynamic Modeling Framework for Hypersonic Aeroelasticity," *51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 2010.

[38] LeGresley, P. A., and Alonso, J. J., "Airfoil Design Optimization Using Reduced Order Models Based on Proper Orthogonal Decomposition," *AIAA Fluids 2000 Conference and Exhibit*, 2000.

[39] Maaten, L. v., Postma, E., and Herik, J. v., "Dimensionality Reduction: A Comparative Review," *Journal of Machine Learning Research*, Vol. 10, No. 66–71, 2009, pp. 1–35.

[40] Tenenbaum, J. B., DeSilva, V., and Langford, J. C., "A Global Geometric Framework for Nonlinear Dimensionality Reduction," *Science*, Vol. 290, No. 5500, 2000, pp. 2319–2323.

[41] Cayton, L., "Algorithms for Manifold Learning," *University of California at San Diego Technical Report*, Vol. 12, No. 1-17, 2005, p. 1.

[42] Bernstein, A. V., and Kuleshov, A. P., "Tangent bundle manifold learning via Grassmann & Stiefel eigenmaps," *arXiv preprint arXiv:1212.6031*, 2012.

[43] Saul, L. K., and Roweis, S. T., "Think Globally, Fit Locally: Unsupervised Learning of Low-Dimensional Manifolds," *Journal of Machine Learning Research*, Vol. 4, 2003, pp. 119–155.

[44] Roweis, S. T., and Saul, L. K., "Nonlinear Dimensionality Reduction by Local Linear Embedding," *Science*, Vol. 290, No. 5500, 2000, pp. 2323–2326.

[45] Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., and Duchesnay, E., "Scikit-learn: Machine Learning in Python," *Journal of Machine Learning Research*, Vol. 12, 2011, pp. 2825–2830.

[46] Shi, L., et al., "A Model of Selecting the Parameters Based on the Variance of Distance Ratios of Manifold Learning Algorithms," *2009 Sixth International Conference on Fuzzy Systems and Knowledge Discovery*, Vol. 2, IEEE, 2009, pp. 507–512.

[47] Virtanen, P., et al., "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python," *Nature Methods*, 2020. doi:https://doi.org/10.1038/s41592-019-0686-2, URL `https://rdcu.be/b08Wh`.

[48] Biedron, R. T., et al., *FUN3D Manual: 13.4*, National Aeronautics and Space Administration, Langley Research Center, Oct. 2018.

[49] Minzer, R. A., Champion, K. S., and Pond, H. L., "The ARDC Model Atmosphere, 1959," *Air Force Surveys in Geophysics*, 1959.

[50] *Release 18.2, Pointwise User Manual*, Pointwise, 2018.

[51] Kholodar, D. B., Morton, S. A., and Cummings, R. M., "Deformation of Unstructured Viscous Grids," *43rd AIAA Aerospace Sciences Meeting and Exhibit*, 2005.

[52] McDaniel, D. R., and Morton, S. A., "Efficient Mesh Deformation for Computational Stability and Control Analyses on Unstructured Viscous Meshes," *47th AIAA Aerospace Sciences Meeting Including the New Horizons Forum and Aerospace Exposition*, 2009.